Using Artificial Intelligence to Automate Body Movement Analysis

Nicholas Connolly and Alireza Goudarzi
University of Iowa Department of Mathematics and UsideU

Exercise Pose Estimation

Computer vision allows for the collection of detailed information about joint motion during an exercise.

- Analyze exercise video frame by frame.
- Apply pose estimation to predict joint positions.
- Extract 3D joint angles from each of 24 joints.
- Each data point corresponds to one video frame: 24 joints $\times$ 3 angles $= 72$ measurements.

Applying this analysis to each frame in the video gives a time series representation of the body's motion.

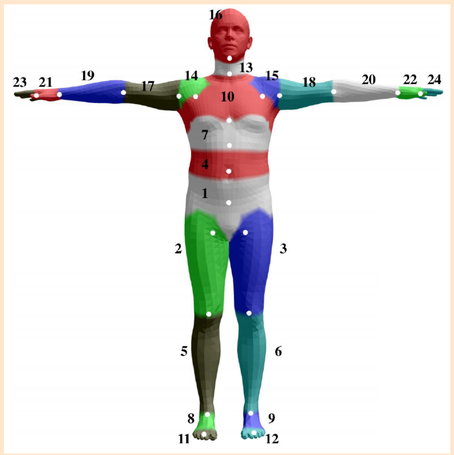


Figure 1: 24 joints of interest

Abstract

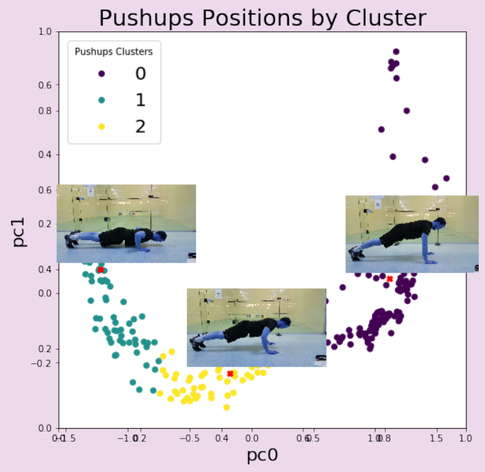
Progress in artificial intelligence has made body movement analysis widely accessible. It is now possible to collect a large amount of detailed data, but processing this data efficiently has become a bottleneck. Traditional rule based systems rely on experts to classify data manually, but this workflow is slow and costly.

Machine learning enables the automation of each task involved with evaluating the performance of an exercise. Data points with similar values can be grouped together with clustering methods, allowing for pose identification without the need for an explicit rule based system. Scoring exercise quality becomes a natural question in deep learning.

Task		Tool
record motion	$\Longleftrightarrow$	pose estimation
identify positions	$\Longleftrightarrow$	joint data clustering
joint selection	$\Longleftrightarrow$	feature contribution
score exercise	$\Longleftrightarrow$	neural network

These advances bypass the costliest steps in the body movement analysis workflow. Presented here is an overview of the exercise evaluation process currently under development by Tokyo-based startup UsideU.

Clustering Positions



Position Sequence:
$[\underbrace{0,2,1,2}_{\text{pushup 1}},\underbrace{0,2,1,2}_{\text{pushup 2}},0]$

Cluster Labels:
1 = down
2 = half pushup
0 = full pushup

Figure 4: Clustered pushups data and position sequence

- Reduce joint angle data to 2 dimensions with PCA.
- Label each video frame using KMeans clustering.
- Cluster labels denote distinct exercise positions.
- Pose duration is tracked in video frames.
- Exercise repetitions encoded by position sequence.

Joint Selection by Regression

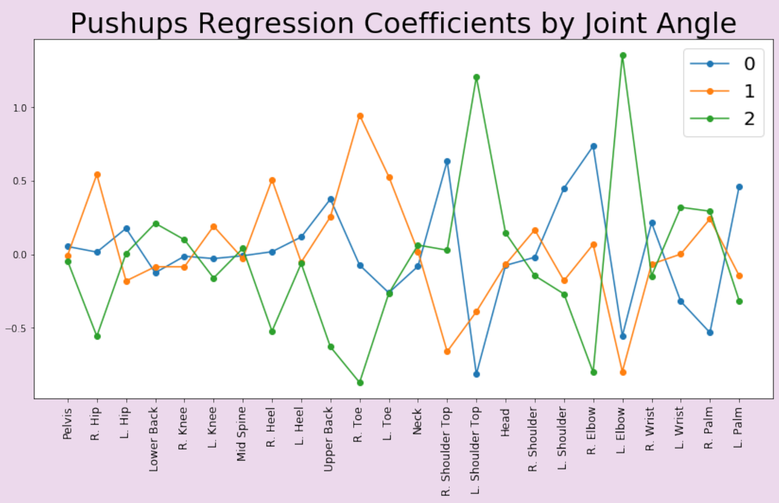


Figure 5: Contribution of each joint to distinguishing pose

- Isolate most important joints with a multi-regression model.
- Features are first component of each triple of joint angles.
- Target label is pose in exercise determined by clustering.

Time Series Joint Data

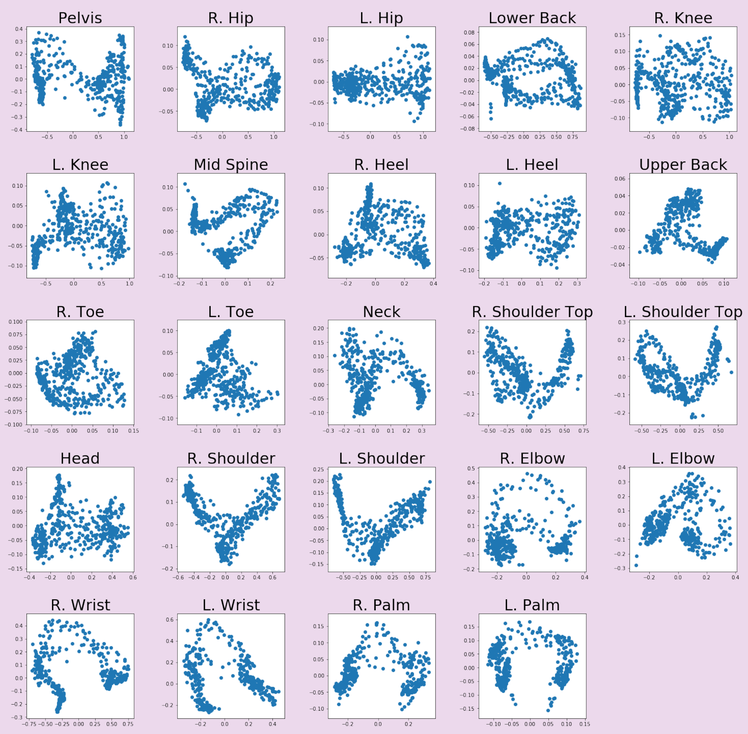


Figure 2: Angular data (first two components) for each of 24 joints from 20 second exercise video

Dimension Reduction

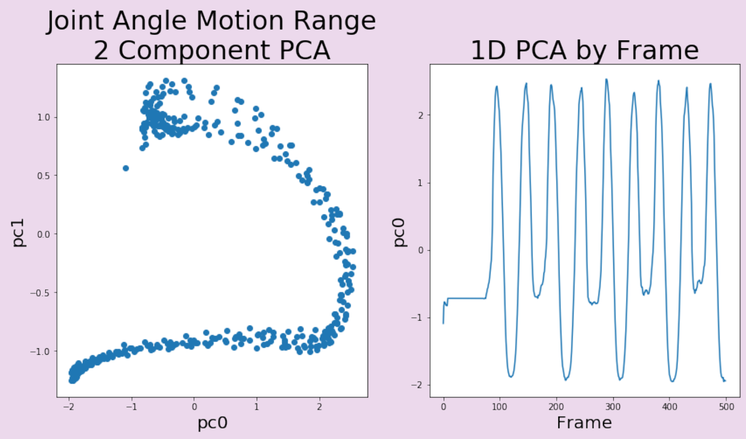
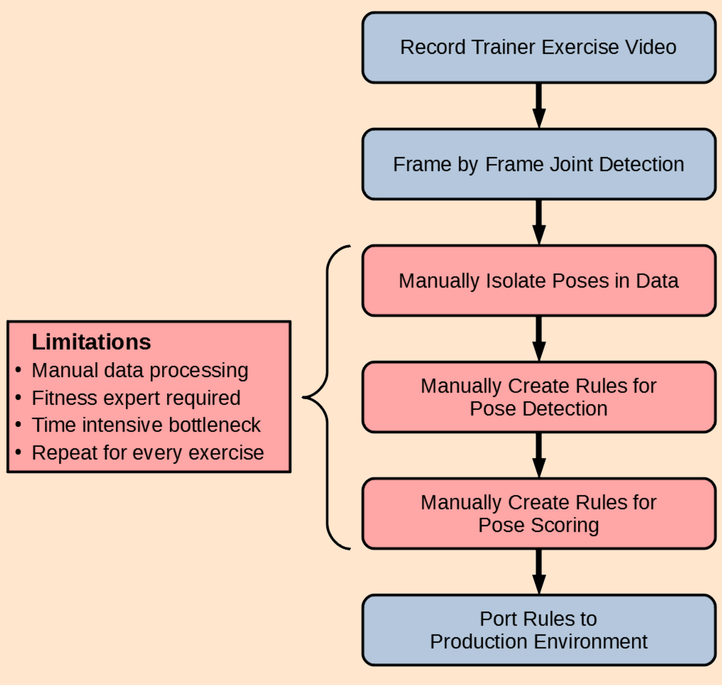


Figure 3: Total angular data reduced with PCA

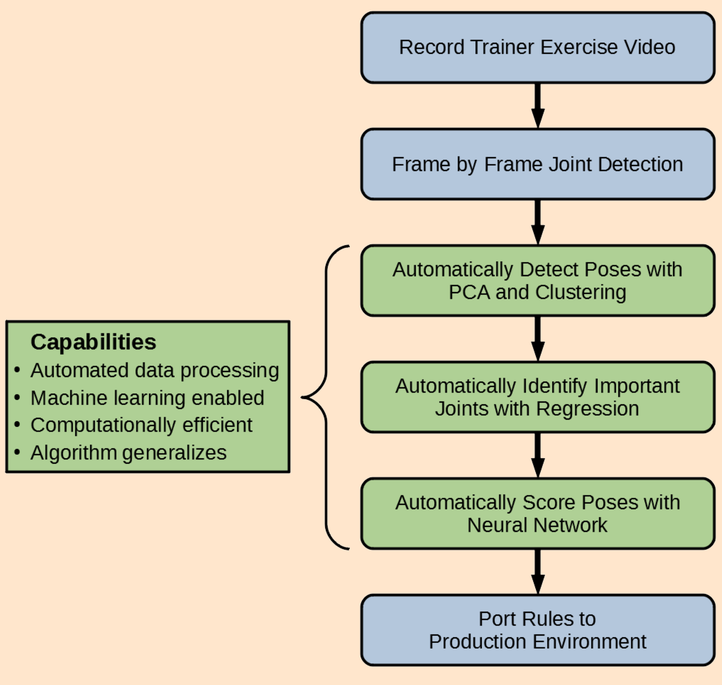
Important features of the joints' motion during an exercise are visible in both the raw data and PCA plots.

- Joint angles follow clear trajectories.
- Measurements oscillate between clusters of data points.
- Clusters of measurements indicate distinct body poses.
- Periodicity in plots reflects repetitions in exercise.
- Repetitions cycle through all poses.

Manual Workflow



Automated Workflow



Neural Network Scoring

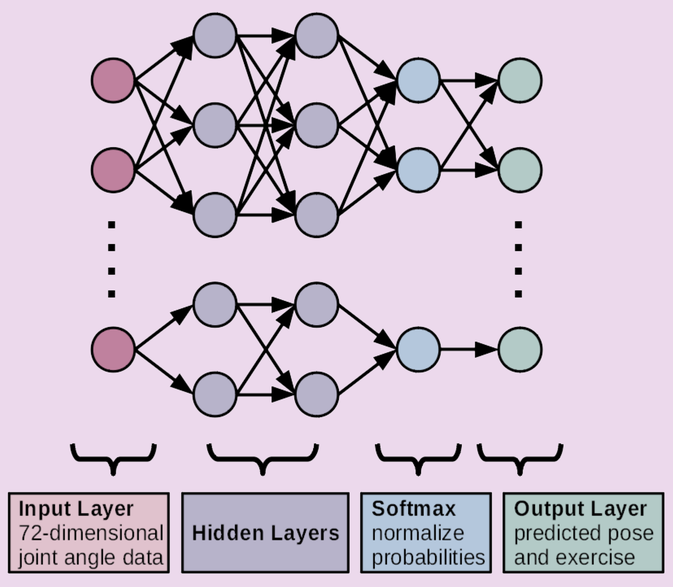


Figure 6: Neural network schematic for exercise scoring

- Train neural network with joint data of fitness experts.
- The input features are the joint data for a video frame.
- The predicted labels include both exercise and position.
- The last hidden layer assigns a probability to each label.

An entire exercise video is scored by averaging the highest probabilities of a subset of video frames nearest the KMeans centroids after clustering.